Estimating a Cholesky Decomposition*

Ingram Olkin
Department of Statistics
Sequoia Hall
Stanford University
Stanford, California 94305

Submitted by George P. H. Styan

ABSTRACT

For a normal distribution the sample covariance matrix S provides an unbiased estimator of the population covariance matrix Σ. We address the problem of finding an unbiased estimator of the lower triangular matrix Ψ defined by the Cholesky decomposition $\Sigma = \Psi\Psi'$.

1. INTRODUCTION

The well-known Cholesky decomposition states that if Σ is a $p \times p$ positive definite matrix that there exists a unique lower triangular matrix Ψ with positive diagonals such that

$$\Sigma = \Psi\Psi'. \tag{1.1}$$

Suppose that we have a "good" estimate, S, of Σ, and carry out a Cholesky decomposition of S:

$$S = TT'. \tag{1.2}$$

Will T also be a good estimate of Ψ? Even when $p = 1$ and we have an unbiased estimate, s^2, of σ^2, we know that s is not an unbiased estimate of σ. However, the bias is a constant and can be removed. Consequently, for $p > 1$

*This work was supported in part by the National Science Foundation Grant MCS 81-04262.

52 Vanderbilt Ave., New York, NY 10017 0024-3795/85/$3.30

we should not expect T to be an unbiased estimator of Ψ. The question we resolve is to find the mean and variance of the elements of T when the underlying model is that of a normal distribution. In this case the elements of T are called rectangular coordinates.

2. DISTRIBUTIONAL RESULTS

Suppose that $x=(x_1,\ldots,x_p)$ has a p-variate normal distribution with mean vector $\mu=(\mu_1,\ldots,\mu_p)$ and positive definite covariance matrix $\Sigma=(\sigma_{ij})$. Based on a sample $(x_{1\alpha},\ldots,x_{p\alpha})$, $\alpha=1,\ldots,N$, of size N, the sample cross-products

$$s_{ij}=\sum_{\alpha}(x_{i\alpha}-\bar{x}_i)(x_{j\alpha}-\bar{x}_j),$$

where $\bar{x}_i=\Sigma_\alpha x_{i\alpha}/N$, $n=N-1$, are unbiased estimates of $n\sigma_{ij}$.

The joint distribution of the elements of $S=(s_{ij})$ is called the Wishart distribution with density function

$$p(S;\Sigma)=c|\Sigma|^{-n/2}|S|^{(n-p-1)/2}\exp\left(-\tfrac{1}{2}\operatorname{tr}\Sigma^{-1}S\right), \tag{2.1}$$

where

$$c^{-1}=2^{np/2}\pi^{p(p-1)/4}\prod_1^p\Gamma\left(\frac{n-i+1}{2}\right).$$

If we define the lower triangular matrix T by (1.2), then the joint distribution of the elements of T is

$$p(T;\Psi)=2^pc\left(\prod_1^p\psi_{ii}^{-n}\right)\left(\prod_1^p t_{ii}^{n-p-1}\right)\left(\prod_1^p t_{ii}^{p-i+1}\right)\exp\left(-\tfrac{1}{2}\operatorname{tr}\Psi^{-1}TT'\Psi'^{-1}\right), \tag{2.2}$$

where the factor $2^p\prod t_{ii}^{p-i+1}$ is the Jacobian of the transformation (1.2). (See e.g. [1].)

If we let

$$U=\Psi^{-1}T, \tag{2.3}$$

then the elements u_{ij} of the lower triangular matrix U will be independently distributed:

$$\mathscr{L}(u_{ij}) \sim \mathscr{N}(0,1), \qquad i > j, \tag{2.4}$$

$$\mathscr{L}(u_{ii}^2) \sim \chi^2_{n-i+1}, \qquad i = 1,\ldots,p. \tag{2.5}$$

It is these latter facts that permit us to readily find the moments of the t_{ij}'s. Note first that $T = \Psi U$, so that

$$t_{ii} = \psi_{ii} u_{ii}, \qquad i = 1,\ldots,p. \tag{2.6}$$

Consequently, we obtain from properties of the chi-square distribution that

$$Et_{ii} = \psi_{ii} E u_{ii} = \psi_{ii} a_i, \tag{2.7}$$

where

$$a_i = \sqrt{2}\,\frac{\Gamma((n-i+2)/2)}{\Gamma((n-i+1)/2)}. \tag{2.8}$$

Again, from (2.3), when $i > j$,

$$t_{ij} = \psi_{ij} u_{jj} + \psi_{i,j+1} u_{j+1,j} + \cdots + \psi_{ii} u_{ij}, \tag{2.9}$$

so that

$$Et_{ij} = a_j \psi_{ij} \tag{2.10}$$

We now incorporate (2.7) and (2.10) into the following result.

THEOREM. *If the sample covariance matrix S has a Wishart distribution with $ES/n = \Sigma$, and we let $\Sigma = \Psi\Psi'$, $S = TT'$, where Ψ and T are lower triangular matrices, then*

$$E(TD_a^{-1}) = \Psi, \tag{2.11}$$

where

$$D_a = \mathrm{diag}(a_1,\ldots,a_p).$$

To obtain the variance of the elements t_{ij}, from (2.6) we have that

$$\operatorname{Var} t_{ii} = 2(n - i + 1)\psi_{ii}^2, \tag{2.12}$$

and from (2.9)

$$\operatorname{Var} t_{ij} = 2(n - j + 1)\psi_{ij}^2 + \left(\psi_{i,j+1}^2 + \cdots + \psi_{ii}^2\right). \tag{2.13}$$

Although TD_a^{-1} is an unbiased estimator of Ψ, there may be a biased estimator with lower mean square error (MSE). Consider the class of estimators TM, where M is lower triangular, with expected loss

$$E\operatorname{tr}(TM - \Psi)(TM - \psi)'. \tag{2.14}$$

This expected loss is to be minimized with respect to M. Let $T = \Psi U$; then (2.14) becomes

$$\begin{aligned} &\operatorname{tr} E\left[\Psi UMM'U'\Psi' - \Psi M'U'\Psi' - \Psi UM\Psi' + \Psi\Psi'\right] \\ &\qquad = \operatorname{tr}\Psi'\Psi\left[E(UMM'U') - M'(EU)' - (EU)M + I\right]. \end{aligned} \tag{2.15}$$

Using (2.4) and (2.5) and the independence of the elements of U, we have that $EU = D_a$, where $a_1, \dots, a_p$ are defined by (2.8). To compute $EUMM'U'$, write $Q = MM'$. The (i, j)th element of UQU' is

$$(u_{i1}, \dots, u_{ii}, 0, \dots, 0)Q(u_{j1}, \dots, u_{jj}, 0, \dots, 0)',$$

which has expectation zero unless $i = j$, in which case we have, as a consequence of (2.4) and (2.5),

$$E\sum_{k,l} u_{ik}q_{kl}u_{il} = \sum_{k,l} q_{kl}Eu_{ik}u_{il} = \sum_{k} q_{kk}Eu_{ik}^2. \tag{2.16}$$

Let $b_i = n - i + 1$, $i = 1, \dots, p$; then (2.15) becomes

$$\operatorname{tr}\Psi'\Psi(D^* - M'D_a - D_aM + I), \tag{2.17}$$

where

$$D^* = \operatorname{diag}\left(b_1q_{11}, b_2q_{22} + q_{11}, \dots, b_pq_{pp} + q_{p-1,p-1} + \cdots + q_{11}\right). \tag{2.18}$$

For (2.17) to be a minimum for all Ψ, M must be diagonal. To see this, let $\Psi'\Psi = \Theta = (\theta_{ij})$.

With $\theta_{11} = 1$ and all other $\theta_{ij} = 0$, (2.17) becomes

$$b_1 q_{11} - 2m_{11}a_1 = b_1 m_{11}^2 - 2m_{11}a_1,$$

which is minimized at $m_{11} = a_1/b_1$. With $\theta_{22} = 1$ and all other $\theta_{ij} = 0$, (2.17) now becomes

$$b_2 m_{12}^2 + \left(b_2 m_{22}^2 - 2m_{22}a_2\right) + \frac{a_1^2}{b_1^2},$$

which is minimized at $m_{12} = 0$ and $m_{22} = a_2/b_2$. This argument can be iterated in a straightforward manner to yield the minimizer

$$M = \operatorname{diag}\left(\frac{a_1}{b_1}, \dots, \frac{a_p}{b_p}\right), \tag{2.19}$$

where

$$\frac{a_i}{b_i} = \sqrt{2}\,\frac{\Gamma((n-i+2)/2)}{(n-i+1)\Gamma((n-i+1)/2)},$$

$i = 1, \dots, p$.

I am indebted to Gene Golub for posing this problem and for subsequent discussions.

REFERENCE

1 W. L. Deemer and I. Olkin, The Jacobians of certain matrix transformations useful in multivariate analysis, *Biometrika* 38:345–367 (1951).